

11/11/11

<http://www1.ipdl.jpo.go.jp/PA1/result/detail/main/wAAAa11224DA413350707P1.h...> 2003/06/06

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-350707

(P2001-350707A)

(43)公開日 平成13年12月21日(2001. 12. 21)

| (51)Int.Cl. ⁷ | 識別記号 | F I | キーワード(参考) |
|--------------------------|-------|---------------|-------------------|
| G 0 6 F 13/14 | 3 1 0 | G 0 6 F 13/14 | 3 1 0 B 5 B 0 1 4 |
| 3/06 | 3 0 1 | 3/06 | 3 0 1 A 5 B 0 6 5 |
| 12/00 | 5 0 1 | 12/00 | 5 0 1 A 5 B 0 8 2 |

審査請求 未請求 請求項の数 7 O L (全 10 頁)

(21)出願番号 特願2000-174203(P2000-174203)

(22)出願日 平成12年6月6日(2000. 6. 6)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 島田 直樹

神奈川県横浜市戸塚区戸塚町5030番地 株
式会社日立製作所ソフトウェア事業部内

(72)発明者 平林 元明

神奈川県横浜市戸塚区戸塚町5030番地 株
式会社日立製作所ソフトウェア事業部内

(74)代理人 100075096

弁理士 作田 康夫

Fターム(参考) 5B014 EB04 FA02 FA05 FB04 HC02

5B065 BA06 CC03 ZA02

5B082 AA01 CA05 EA07 GA20 HA00

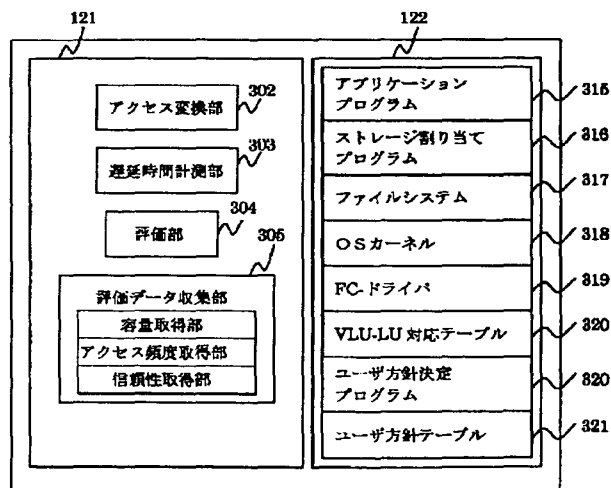
(54)【発明の名称】 情報処理システム、記憶装置の割り当て方法

(57)【要約】

【課題】ストレージ装置あるいはストレージ装置の論理ユニット単位で割り当てを可能とするシステムを提供する。

【解決手段】ホスト計算機111内で仮想的に設定された仮想ロジカルユニット(VLU)とストレージ装置の論理ユニット(LU)とを対応付けするVLU-LU対応テーブル320を設け、このテーブルに遅延時間、アクセス頻度などの評価項目を格納する。評価部304では、評価項目から評価値を求め、この評価値に基づいて割り当てを行う。

図 3



【特許請求の範囲】

【請求項 1】 ネットワークに接続された複数の記憶装置における所定の記憶領域を、前記ネットワークに接続された計算機に割り当てる記憶装置の割り当て方法であって、

割り当て可能な所定の記憶領域に関する情報を収集し、前記収集した情報に基づいて前記所定の記憶領域ごとの評価値を求め、当該求めた評価値から割り当てるべき所定の記憶領域を決定する記憶装置の割り当て方法。

【請求項 2】 請求項 1 記載の記憶装置割り当て方法において、

前記所定の記憶領域に関する情報は、前記所定の記憶領域を有する記憶装置のアクセス頻度、容量、遅延時間である記憶装置の割り当て方法。

【請求項 3】 ネットワークに接続された複数の記憶装置における所定の記憶領域を、前記ネットワークに接続された複数の計算機に割り当てる記憶装置の割り当て方法であって、

割り当て可能な所定の記憶領域に関する情報を収集し、前記収集した情報に基づいて前記所定の記憶領域ごとの評価値を求め、当該求めた評価値と予め決定された条件に基づいて割り当てるべき所定の記憶領域を決定する記憶装置の割り当て方法。

【請求項 4】 ホスト計算機と複数の記憶装置がネットワークに接続された情報処理システムであって、前記ホスト計算機は、前記記憶装置の所定の領域に関する情報を記憶する記憶部と、アプリケーションプログラムの実行により該所定の領域に関する情報に基づいて該所定の領域毎の評価値を求め、該評価値に基づいて割り当てるべき所定の記憶領域を決定する処理部とを有する情報処理システム。

【請求項 5】 複数のホスト計算機と複数の記憶装置がネットワークに接続された情報処理システムであって、それぞれの前記記憶装置は、それぞれの前記ホスト計算機と該ホスト計算機に対応付けされた前記記憶装置の所定の記憶領域と、該所定の記憶領域に関する情報を格納する記憶部を有し、前記ホスト計算機は、アプリケーションプログラムの実行により記憶装置を選択し、選択した記憶装置の記憶部に格納されている前記所定の記憶領域に関する情報に基づいて前記所定の記憶領域毎に評価値を求め、該評価値に基づいて割り当てるべき所定の記憶領域を決定する処理部とを有する情報処理システム。

【請求項 6】 請求項 4 又は 5 記載の情報処理システムにおいて、前記所定の記憶領域に関する情報は、前記所定の記憶領域を有する記憶装置のアクセス頻度、容量、遅延時間である情報処理システム。

【請求項 7】 ネットワークに接続された複数の記憶装置における所定の記憶領域を、前記ネットワークに接続さ

れた計算機に割り当てる処理であって、

割り当て可能な所定の記憶領域に関する情報を収集する処理と、前記収集した情報に基づいて前記所定の記憶領域ごとの評価値を求め、当該求めた評価値から割り当てるべき所定の記憶領域を決定する処理とを記憶したコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明はネットワークに接続されたストレージ装置の割り当てに関するものである。

【0002】

【従来の技術】 特開平 10-69357 号公報にネットワークで相互に接続されたディスク制御装置について開示されている。ここでは、複数のディスクアレイ制御装置をネットワークで接続した形態のストレージシステムにおいて、ホストとディスクアレイ制御装置間における通信プロトコルと、ディスクアレイ制御装置間における通信プロトコルとを切り替える機構を設けることが開示されている。これにより、ホストとは汎用インターフェースを保ち、ディスクアレイ制御装置間は高速に通信を行うようにしたものである。

【0003】 また、特開 2000-99272 号公報も同様にネットワークで相互に接続されたディスク制御装置について開示されている。ここでは、ファイバチャネルネットワークに接続された既存の記憶制御装置に対し、ファイバチャネルネットワーク上に新たな記憶制御装置をオンライン中に追加し、ロジカルユニットの処理担当を引き継がせることが開示されている。これにより、記憶制御装置間の負荷分散、処理の統合を可能にして、システム全体の処理能力を向上させるようにしたものである。

【0004】

【発明が解決しようとする課題】 ネットワークにホスト計算機とストレージ装置が接続されたシステムにおいて、従来はホスト計算機が利用するストレージ装置あるいはロジカルユニットは予め定められているものであった。

【0005】 ネットワークに複数のストレージ装置が接続されたストレージシステムにおいて、各ストレージ装置はネットワークに接続された複数のホスト計算機でアクセスが可能になる。また、ストレージ装置間においてもアクセスが可能になる。

【0006】 このように構成されたシステムにおいて、ホスト計算機がストレージにアクセスする場合、ネットワークの伝送路によりストレージ装置毎にアクセスに要する時間が異なる場合がある。また、ストレージ装置毎にアクセス頻度やバックアップの頻度が異なる場合があり、ストレージ装置毎に信頼性が異なる場合がある。

【0007】 また、ホスト計算機においてもアプリケーションプログラムによってはアクセスに要する時間が早

いものが必要であったり、アクセスに要する時間が遅くても信頼性が必要であったりする。

【0008】従って、アプリケーションプログラム毎にストレージ装置あるいは論理ユニットの割り当てを行えるようにすれば、より効率の良いシステムを構築することが可能となる。

【0009】本発明の目的は、ホスト計算機側でストレージ装置あるいは論理ユニットの割り当てを行うことを可能とする方法及びシステムを提供することにある。

【0010】

【課題を解決するための手段】上記目的を達成するために、本発明では割り当てられる所定の記憶領域毎に、この記憶領域に関する情報を管理し、割り当てを行う際に、この情報に基づいて記憶領域毎に評価値を求め、この評価値に基づいて割り当てるべき記憶領域を決定する。

【0011】これにより、記憶装置毎にあるいは記憶領域毎に割り当てを行うことが可能となり、効率の良いシステムを構築することが可能となる。

【0012】

【発明の実施の形態】本発明の一実施例を、以下図面に従って説明する。

【0013】図1は本発明によるストレージの割当て方法を適用したシステムの全体構成を示したものである。ホスト計算機111は、Fabric Channel スイッチ113（以下、「FC-Switch 113」と称す。）を介して複数のストレージ装置117、118と接続されている。更にストレージ装置116、119はFC-Switch 115を介して接続されている。ストレージ装置116、117、118、119は光ディスク装置、磁気ディスク装置、光磁気ディスク装置あるいは磁気テープのいずれであっても良い。

【0014】ストレージ装置116、117、118、119は、それぞれディスクとディスクコントローラを備えている。図2はストレージ装置のディスクコントローラの構成を示したものである。ディスクコントローラは、CPUなどによって構成された処理部211、RAMやROMなどによって構成された記憶部212を有している。記憶部212にはFC-Switchやネットワークを制御するためのネットワーク処理プログラム、ストレージ装置の制御プログラムが格納されており、処理部211によって、このプログラムを起動することによりFC-Switch 113、115を制御し、ホスト計算機111を介さずにストレージ装置116、117、118、119間だけで直接データのやり取りが行えるようになってい

る。尚、このホスト計算機111は別のネットワークを

介して上位の計算機と接続される構成であっても良い。

【0016】このようなシステムにおいて、ホスト計算機111の処理部121は、アプリケーションプログラムの実行によりストレージ装置にアクセスする必要がある場合には、ストレージ割り当てプログラムを実行し、予めユーザから与えられている方針テーブルに基づいて、ユーザから与えられている方針（条件）を満たすあるいは方針（条件）に近いストレージ装置を決定し、このストレージ装置のファイルやデータベースを利用することができる。従って、例えばユーザが「性能（アクセス速度）」といった方針（条件）をホスト計算機111に与えた場合、1つのFC-Switch 113を介してアクセス可能なストレージ装置117又はストレージ装置118が決定される。

【0017】本実施例において、ホスト計算機111は仮想的な記憶領域を有しており、ネットワーク内で接続許可されたストレージ装置を可能なだけ利用することができる。ホスト計算機111が有している仮想的な記憶領域は、ある記憶領域毎に仮想ロジカルユニット（VLU）として仮想ロジカルユニット番号（VLUN）を付して管理している。また、各ディスク装置のディスク毎などのように所定の記憶領域の単位で論理ユニット（LU）が設けられ、それぞれ論理ユニット番号（LUN）を付して管理している。図10は、仮想ロジカルユニット（VLU）と論理ユニット（LU）との対応関係を示したものである。図では仮想ロジカル論理ユニット毎に別々のストレージB1002、D1003の論理ユニット（LU）を指定し、あたかも連続領域であるかのような仮想ロジカル論理ユニット（VLUN）が二つ存在している状態を示している。

【0018】図3はホスト計算機111の構成を示したものである。

【0019】ホスト計算機111の記憶部122は、アプリケーションプログラム315、ストレージ割り当てプログラム316、ファイルの管理を行うためのファイルシステム317、メモリ管理やタスク管理を行うOSカーネル318、FC-SWITCHを駆動するFCードライバ319、それぞれのディスク装置の遅延時間やアクセス頻度等が記録されているVLU-LU対応テーブル320、ユーザの方針を入力するユーザ方針決定プログラム321、ユーザ方針決定プログラム321で決定された方針を記憶する方針テーブル322とが保持されている。ここで、アプリケーションプログラム315、ストレージ割り当てプログラム316、ユーザの方針を入力するユーザ方針決定プログラム321は、予めホスト計算機111の記憶部121へ記憶されている場合であっても良いし、磁気ディスク装置などのコンピュータで読み取りが可能な記憶媒体に格納されており、この記憶媒体から読み出して、ホスト計算機111の記憶部122へ格納されたものでもよい。

【0020】処理部121は、ストレージ割り当てプロ

グラムを実行するための構成を示している。処理部121は、仮想論理ユニット番号(VLUN)と論理ユニット番号(LUN)との変換を行うアクセス変換部302、ホスト計算機111からストレージ装置の論理ユニット(LU)の応答時間を遅延時間として計測する遅延時間計測部303、ホスト計算機111がアクセスするストレージ装置の論理ユニット(LU)を決定する評価部304、評価部304で評価するための項目値を収集する評価データ収集部305とを有している。

【0021】評価データ収集部305は、OSのコマンドを用いて各ストレージ装置の空き容量を計測する容量取得部、各ストレージ装置へのOpen/Create/Read/Write命令があった回数を取得するアクセス頻度取得部、ストレージ装置が他のデバイスにバックアップされている頻度を取得したり、ストレージ装置にボリューム二重書きの機能がありそれを利用しているか否かを調べる信頼性取得部とを有している。

【0022】尚、ここで遅延時間計測部303を評価データ収集部305と別にしてしているのは、遅延時間計測部204はストレージ割当ての要求があった時に動作し、評価データ収集部305は定期的に動作するようにしているためである。これにより、遅延時間はその時点において可能な限り正確な情報とすることができる。

【0023】図4はVLUN-LU対応テーブルの一例を示したものである。キー項目となるのは仮想ロジカルユニット(VLU)毎にユニークな番号である仮想ロジカルユニット番号(VLUN)402と物理的な位置を示す絶対セクタ番号403である。論理ユニット(LU)の位置を特定するために同様に論理ユニット番号(LU)404と物理的な位置を示す絶対セクタ番号405を用意する。仮想ロジカルユニット番号(VLUN)に論理ユニット番号(LUN)が対応づけられていないとき仮想ロジカルユニット番号(VLUN)の値は0に設定されている。絶対セクタ番号は等容量ごとに割り振られており論理ユニット(LU)の個数を乗じた値が仮想ロジカルユニット(VLU)の容量を示している。

【0024】遅延時間406はストレージ装置への割当て要求のあったときに、遅延時間計測部303によって書き換えられる項目である。尚、この値はユーザの特殊な要求がない限り小さいほど良いものとする。

【0025】アクセス頻度407は、それぞれの論理ユニット(LU)への参照・書き込みの回数を示しており、性能や信頼性に影響を与える。その時点で未使用のLUの場合は0となる。この値も小さいほど良いものとする。

【0026】バックアップ頻度408は、それぞれの論理ユニット(LU)が他のデバイスにバックアップを取られている年当たりの回数を示している。バックアップが取られていない場合はこの値が0となる。この値は大きいほど信頼性も良いと判断する。

【0027】信頼性を高めるため、ストレージの構成に

よっては同一内容を複数のボリュームに書き込む場合がある。この機能を使用しているかどうかはボリュームの二重書き有無409である。当然使用している方が信頼性も高いとみなす。

【0028】以上の評価項目の他、更に小さい単位でのディスクの冗長化の程度や、遅延時間の統計的推移を採取することでストレージ割当ての指針のバリエーションを豊富にすることができる。

【0029】評価値410は評価部304で求められた各論理ユニット(LU)の評価値を示したものである。この評価値に基づいて仮想ロジカルユニット(VLU)と論理ユニット(LU)との対応づけが行われる。尚、仮想ロジカルユニット番号(VLUN)に正の値が入っている場合は以前の評価値が残っており、0の場合はその時点での評価値を示している。対応づけされている論理ユニット(LU)の評価値で対応づけされていない論理ユニット(LU)の評価値よりも小さい値しか持たないものもあるが、これは評価された時点が異なるためか、あるいはユーザの方針が異なっていたために生じる現象である。

【0030】以上の構成において、ユーザの方針の決定およびストレージ装置の割り当て動作を以下説明する。

【0031】まず、ユーザの方針を決定する処理について説明する。

【0032】ユーザの方針の決定は、ユーザ方針決定プログラム320を処理部121で実行することにより行われる。図5はユーザ方針決定プログラム320の処理フローを示したものである。

【0033】まず、重視事項において評価項目を抽象化した特性を表示する(501)。図6は、この重視事項の表示例とそれに対するユーザ入力の例を示したものである。ユーザ方針決定プログラム320はユーザの入力を受け付ける(502)。本実施例では図6に示すように重視事項601として評価項目を抽象化した特性を並べて表示し、ユーザがこの重視事項を選択できる構成となっている。本例では、ユーザによって信頼性と容量が選択されている。

【0034】この重視事項の各特性とVLUN-LU対応テーブルの評価項目とは図7に示すテーブルによって対応づけされている。この図7に示すテーブルはユーザが書き換えることが可能である。

【0035】また、ユーザによって選択された重視事項に対しては、遵守するか(必須607)、可能な限り希望に沿うか(優先609)の2通りを厳密性606として選択することが可能になっている。

【0036】「優先」を選んだ場合には、仮想ロジカルユニット(VLU)と対応付けされていない論理ユニット(LU)全てについて評価値を求め、「必須」を選んだ場合には閾値を満たす論理ユニット(LU)について評価値を求める。閾値として例えば、バックアップ頻度が52回/年以上、かつボリューム二重書き有無が1の論理ユ

ニット (LU) 等があり、ユーザはこの閾値を指定することが可能となっている。

【0037】ユーザによって選択された重視事項に基づいて、対応する評価項目の評価係数を変換する(503)。例えばGUIの重視事項501で信頼性を選択した場合、アクセス頻度の評価係数を1.5倍にし、バックアップ頻度とボリューム二重書き有無の評価係数を2倍にする。

【0038】このようにして変更された評価係数及び「優先」、「厳密」についてユーザ方針テーブル321へ書き込む。

【0039】尚、図8はユーザがVLU-LU対応テーブルを参照する場合の表示例を示したものである。ユーザはこのテーブルを見ながら重視事項601や厳密性606を決定したり、現在の対応づけや未対応のLUを確認する。表示項目の設定802ではVLU-LU対応テーブルの表示項目を増減することができ、デフォルトである全てを表示した場合は801のようになる。

【0040】次にストレージの割り当て処理について説明する。

【0041】ストレージの割り当てはストレージ割り当てプログラム316を実行することにより行われる。

【0042】図9はストレージ割り当てプログラムの処理フローを示したものである。

【0043】ホスト計算機111の処理部121でアプリケーションプログラム315が起動され、ファイル名を指定しファイルのOpen/Create命令が発行されると、この命令はファイルシステム317へ送られる。ファイルシステム317はディレクトリ構造を参照しファイルIDをアプリケーションプログラム315へ返す。アプリケーションプログラム315はファイルのRead/Writeのために、この受け取ったファイルIDと相対セクタ番号をファイルシステム317に送付する。ファイルシステム317では、これらから絶対セクタ番号を割り出し、ファイルIDに含まれるドライブ番号と共にOSのカーネル318に送付する。OSカーネル318ではドライブ番号を仮想ロジカルユニット番号(VLUN)に変換し、ストレージ割り当てプログラム316を起動し、絶対セクタ番号と共に送付する。

【0044】ストレージ割り当てプログラムが起動されるとアクセス変換部202ではVLU-LU対応テーブルを参照し(901)、当該仮想ロジックユニット(VLU)があれば対応する論理ユニット番号(LUN)を戻す(902)。なければ新たな論理ユニット(LU)を割り当てることになるが、まずユーザ方針テーブル、VLU-LU対応テーブルを参照して評価値を算出する論理ユニット(LU)を検索する(903)。論理ユニット(LU)の検索は、ユーザ方針テーブルで「必須」となっている評価項目がある場合、設定された閾値を満足する論理ユニット(LU)を抽出する。ここで、全ての論理ユニットが満足し

ない場合には、「閾値該当論理ユニット(LU)無し。

「優先」処理を実行します。」と表示し、ユーザに「優先」で処理することを知らせ、ユーザ方針テーブルの「必須」を「優先」に変える。ユーザ方針テーブルに、他の「必須」がある場合は同様に閾値を満足する論理ユニット(LU)を抽出する。

【0045】一方、「必須」がない場合には仮想ロジカルユニット(VLU)に対応付けされていない論理ユニット(LU)を抽出する。

【0046】次に抽出された論理ユニット(LU)に対して遅延時間計測部303により、遅延時間の計測を行う(904)。遅延時間の計測は、タイマーをスタートさせ、ファイルシステムの種類に応じてマウント要求ないし参照要求を発行し、応答を受けるまでの時間を計測することにより行う。

【0047】遅延時間計測部303により測定された結果をVLU-LU対応テーブルに書き込む(905)。

【0048】次に、評価部304でユーザ方針テーブルから評価係数を読み出してVLU-LU対応テーブルの各評価項目の重み付けを図り、各評価項目値を $S=K1*H1+K2*H2 \dots$ (Sは評価値、Knは評価係数、Hnは評価項目値、)の多項式に展開する(906)。値の小さいほど良いとされる評価項目値は逆数に変換する。又、各評価項目値間でバランスが取れるよう適切な評価係数を設定する。例えば遅延時間はミリ秒を単位としているが他の評価項目値に比べ全般的に値が小さくなるため、評価係数を2倍にしておく。

【0049】全ての論理ユニット(LU)に関し評価値が求まると、その値の高い順に新規に作成するVLUの値と対応づけるべくVLU-LU対応テーブルの仮想ロジックユニット番号(VLUN)の項目402を埋めてゆく(907)。

【0050】アプリケーションからの要求に容量が含まれている場合は同一VLUNを持つLUの個数を増減して対応する。

【0051】以上説明したように本発明によれば、ユーザの割り当て方針に従って、ストレージ装置を割り当てることが可能となる。

【0052】尚、本実施例では説明していなかったが、VLU-LU対応テーブルの評価項目としてストレージ装置の空き容量を追加することも可能である。これにより、空き容量の多いストレージ装置の論理ユニットを選択することも可能になる。

【0053】また、ストレージ装置毎に論理ユニットを割り当てるようにすれば、ストレージ割り当てプログラムは、ストレージ自身の割り当てを行うようになる。

【0054】つまり、本実施例によればVLU-LU対応テーブルの評価項目と図7に示したテーブルを書き換えることにより、さまざまな形態で記憶領域を割り当てることが可能となる。

【0055】次に、他の実施例を説明する。

【0056】図11は本発明によるストレージの割り当て方法を適用したシステムの全体構成を示したものである。

【0057】本システムは、ネットワークに接続された複数のホスト計算機1101、1102、1103と、ネットワークに接続された複数のストレージ装置1111、1112から構成されている。図1で示したものと同一符号が付されているものは図1と同様の機能を備えている（以下の説明においても同じ）。ストレージ装置1111、1112は、それぞれディスクコントローラと複数のディスク装置から構成されている。

【0058】各ホスト計算機1101、1102、1103の構成は図3で示した構成とほぼ同様である。

【0059】但し、本実施例においてストレージ割り当てプログラム316の処理内容は相違するので、後で詳細に説明する。

【0060】また、本実施例において、VLU-LU対応テーブル320は、図4に示したもののよりも簡易なものであってよい。つまり、少なくとも仮想ロジカルユニット番号（VLUN）、仮想ロジカルユニットの絶対セクタ番号、論理ユニット番号（LUN）、論理ユニットの絶対セクタ番号を備えていれば良い。また、このVLU-LU対応テーブルは、後に説明するように仮想ロジカルユニット（VLU）と論理ユニット（LU）が対応付けされたものだけが保持されている。

【0061】図12は、ストレージ装置1111、1112のディスクコントローラを示したものである。ディスクコントローラはCPUなどによって構成された処理部211とRAMやROMで構成された記憶部212から構成されている。記憶部212はネットワークアプリケーションプログラム、制御プログラムの他に、マシン-LU対応テーブル1201が格納されている。

【0062】図13はマシン-LU対応テーブルの一例を示したものである。マシン-LU対応テーブルのキー項目となるのは各ホスト計算機を示すマシン識別子1302と物理的な位置を示す絶対セクタ番号1303である。論理ユニット（LU）の位置を特定するために同様に論理ユニット番号（LU）1304と物理的な位置を示す絶対セクタ番号1305を用意する。マシン識別子に論理ユニット番号（LUN）が対応づけられていないときマシン識別子の値は0に設定されている。絶対セクタ番号は等容量ごとに割り振られており論理ユニット（LU）の個数を乗じた値がホスト計算機の仮想ロジカルユニット（VLU）の容量を示している。

【0063】遅延時間1306はストレージ装置への割り当て要求のあったときに、遅延時間計測部303によって書き換えられる項目である。

【0064】アクセス頻度1307は、それぞれの論理ユニット（LU）への参照・書き込みの回数を示してお

り、性能や信頼性に影響を与える。その時点で未使用のLUの場合は0となる。

【0065】バックアップ頻度1308は、それぞれの論理ユニット（LU）が他のデバイスにバックアップを取られている年当たりの回数を示している。バックアップが取られていない場合はこの値が0となる。信頼性を高めるため、ストレージの構成によっては同一内容を複数のボリュームに書き込む場合がある。この機能を使用しているかどうかはボリュームの二重書き有無1309である。

【0066】以上の評価項目の他、更に小さい単位でのディスクの冗長化の程度や、遅延時間の統計的推移を採取することでストレージ割当ての指針のバリエーションを豊富にすることができる。

【0067】評価値1310は評価部304で求められた各論理ユニット（LU）の評価値を示したものである。この評価値に基づいてホスト計算機と論理ユニット（LU）との対応づけが行われる。尚、マシン識別子に正の値が入っている場合は以前の評価値が残っており、0の場合はその時点での評価値を示している。対応づけされている論理ユニット（LU）の評価値で対応づけされていない論理ユニット（LU）の評価値よりも小さい値しか持たないものもあるが、これは評価された時点が異なるためか、あるいはユーザの方針が異なっていたために生じる現象である。

【0068】次に、ホスト計算機によるストレージ装置の割り当て処理について説明する。尚、ユーザの方針の決定については図5で説明したのと同様の処理によって行われる。

【0069】ストレージ装置の割り当て処理は、全てのホスト計算機で同様に行われる。ここでは、一例としてホスト計算機1101がストレージ割り当て処理を実行する場合について説明する。

【0070】図14は本実施例におけるストレージ割り当てプログラムの処理フローを示したものである。

【0071】ホスト計算機1101の処理部121でアプリケーションプログラム315が起動され、ファイル名を指定しファイルのOpen/Create命令が発行されると、この命令はファイルシステム317へ送られる。ファイルシステム317はディレクトリ構造を参照しファイルIDをアプリケーションプログラム315へ返す。アプリケーションプログラム315はファイルのRead/Writeのために、この受け取ったファイルIDと相対セクタ番号をファイルシステム317に送付する。ファイルシステム317では、これらから絶対セクタ番号を割り出し、ファイルIDに含まれるドライブ番号と共にOSのカーネル318に送付する。OSカーネル318ではドライブ番号を仮想ロジカルユニット番号（VLUN）に変換し、ストレージ割り当てプログラム316を起動し、絶対セクタ番号と共に送付する。

【0072】ストレージ割り当てプログラム316が起動されるとアクセス変換部202ではVLU-LU対応テーブルを参照し(1401)、当該仮想ロジックユニット(VLU)があれば対応する論理ユニット番号(LUN)を戻す(1402)。なければ新たな論理ユニット(LU)を割り当てることになるが、この処理は各ストレージ装置毎に行う。

【0073】まず、ストレージ装置1111のアクセス排他権を取得し、ストレージ装置1111のマシン-LU対応テーブルを読み出す(1403)。次にユーザ方針テーブルを参照して評価値を算出する論理ユニット(LU)を検索する(1404)。論理ユニット(LU)の検索は、ユーザ方針テーブルで「必須」となっている評価項目がある場合、設定された閾値を満足する論理ユニット(LU)を抽出する。ここで、全ての論理ユニットが満足しない場合には、「閾値該当論理ユニット(LU)無し。」「優先」処理を実行します。」と表示し、ユーザに「優先」で処理することを知らせ、ユーザ方針テーブルの「必須」を「優先」に変える。ユーザ方針テーブルに、他の「必須」がある場合は同様に閾値を満足する論理ユニット(LU)を抽出する。

【0074】一方、「必須」がない場合にはマシン識別子に対応付けされていない論理ユニット(LU)を抽出する。

【0075】次に抽出された論理ユニット(LU)に対して遅延時間計測部303により、遅延時間の計測を行う(1405)。遅延時間の計測は、既に説明したのと同様に行う。

【0076】遅延時間計測部1405により測定された結果をマシン-LU対応テーブルに書き込む(1406)。

【0077】次に、評価部304でユーザ方針テーブルから評価係数を読み出してマシン-LU対応テーブルの各評価項目の重み付けを図り、各評価項目値を $S=K1*H1+K2*H2\cdots$ (Sは評価値、Knは評価係数、Hnは評価項目値、)の多項式に展開する(1407)。

【0078】全ての論理ユニット(LU)に関し評価値が求まると、その値の高い論理ユニット番号を評価値と共に記憶する(1408)。

【0079】次に、ストレージ装置1112のアクセス排他権を獲得し、以上説明した1403~1408の処理を実行する。

【0080】このようにして各ストレージ装置から評価値の高い論理ユニット番号が抽出されたので、次にこの抽出された評価値の中で高い評価値の論理ユニットとホスト計算機1101の仮想ロジカルユニットとを対応付ける(1409)。

【0081】この対応付けを行った仮想ロジカルユニット番号(VLUN)、仮想ロジックユニットの絶対セクタ、論理ユニット番号(LUN)、論理ユニットの絶対セクタとをVLU-LU対応テーブルに格納する(1410)。更

に、対応付けされた論理ユニットのストレージ装置へ論理ユニット番号とホスト計算機1111のマシン識別子を送る(1411)。ストレージ装置の処理部は送られてきた論理ユニット番号とホスト計算機1111のマシン識別子からマシン-LU対応テーブルの該当部分にマシン識別子を書き込む。

【0082】アプリケーションからの要求に容量が含まれている場合は同一マシン識別子を持つLUの個数を増減して対応する。

【0083】以上説明したように本発明によれば、ユーザの割り当て方針に従って、ストレージ装置を割り当てるのが可能となる。またマシン-LU対応テーブルをストレージ側に用意し、マシン間で排他的に共有することで同一ストレージに対する割当ての競合を防ぐことができる。

【0084】

【発明の効果】本発明によれば、きめ細かくストレージ装置あるいは記憶領域を割り当てるのが可能となりシステムの効率を向上させることが可能となる。

【図面の簡単な説明】

【図1】本発明によるストレージの割当て方法を適用したシステム全体構成例を示した図である。

【図2】ディスクコントローラの構成を示した図である。

【図3】ホスト計算機の構成を示した図である。

【図4】VLU-LU対応テーブルの例を示した図である。

【図5】ユーザ方針決定処理の処理フローを示した図である。

【図6】画面表示例を示した図である。

【図7】重視事項と評価項目とを対応づけるテーブルを示した図である。

【図8】画面表示例を示した図である。

【図9】ストレージ割り当て処理の処理フローを示した図である。

【図10】仮想ロジカル論理ユニットと論理ユニットとの対応関係を示した図である。

【図11】本発明を適用した他のシステムの構成を示した図である。

【図12】ディスクコントローラの構成を示した図である。

【図13】マシン-LU対応テーブルの一例を示した図である。

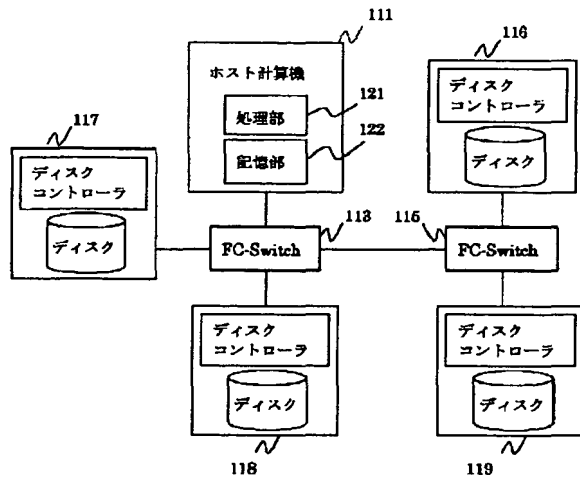
【図14】ストレージ割り当て処理の処理フローを示した図である。

【符号の説明】

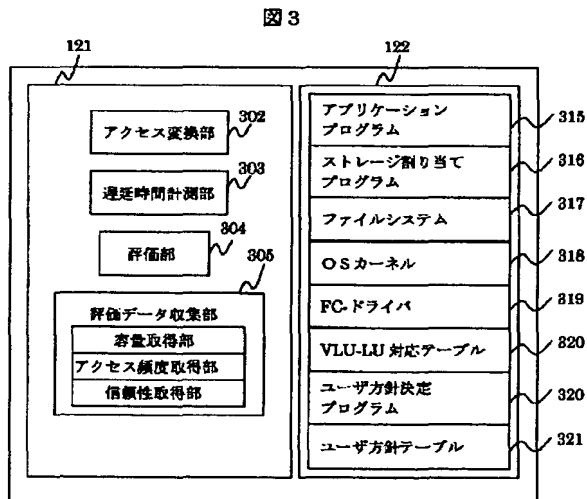
111…ホスト計算機、113、115…FC-Switch、116、117、118、119…ストレージ装置、121…処理部、122…記憶部、211…処理部、212…記憶部、302…アクセス変換部、303…遅延時間計測部、304…評価部、305…評価データ収集部

【図1】

図1



【図3】



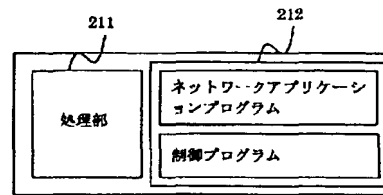
【図7】

図7

| | |
|-----|-----------------------------|
| 信頼性 | アクセス頻度、バックアップ頻度、ボリューム二重書き有無 |
| 性能 | 遅延時間 |
| 信頼度 | アクセス頻度 |
| 容量 | |

【図2】

図2



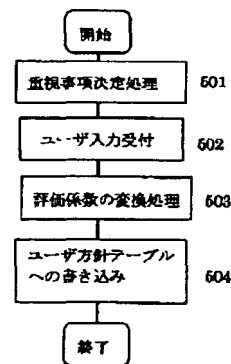
【図4】

図4

| 401 | 402 | 403 | 404 | 405 | 406 | 407 | 408 | 409 | 410 |
|--------|------|-----|--------|------|---------|--------|---------|-----|------|
| VLUN | 絶対値 | LUN | 絶対値 | 遅延時間 | アクセス | バックアップ | ボリューム二重 | | 評価値 |
| 番号(MF) | | | 番号(MH) | (ms) | 頻度(k/日) | 頻度(年) | 書き有無 | | |
| 1 | 0- | 1 | 0- | 57 | 324 | 365 | 0 | | 923 |
| 1 | 100- | 2 | 100- | 42 | 163 | 365 | 0 | | 1186 |
| 1 | 200- | 3 | 0- | 147 | 0 | 52 | 0 | | 1036 |
| 2 | 0- | 11 | 200- | 174 | 122 | 52 | 1 | | 1216 |
| 2 | 100- | 12 | 0- | 33 | 273 | 52 | 1 | | 1337 |
| 0 | 0- | 6 | 100- | 67 | 142 | 52 | 0 | | 1079 |
| 0 | 100- | 7 | 100- | 45 | 79 | 365 | 0 | | 1369 |
| 0 | 0- | 24 | 0- | 37 | 115 | 52 | 1 | | 1243 |

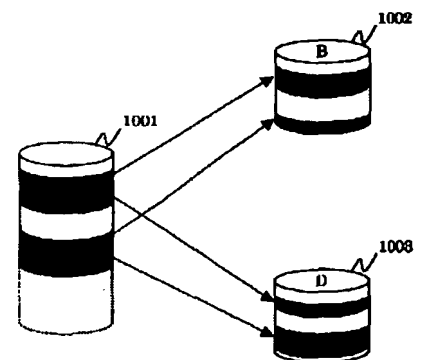
【図5】

図5



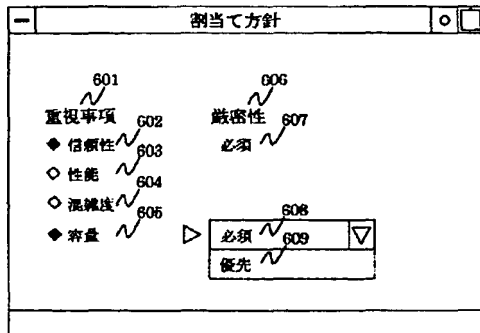
【図10】

図10



【図6】

図6



【図8】

図8

VLU-LU 対応テーブル

表示

項目設定

VLU 番号

VLU 絶対セクタ番号

LU 番号

LU 絶対セクタ番号

...

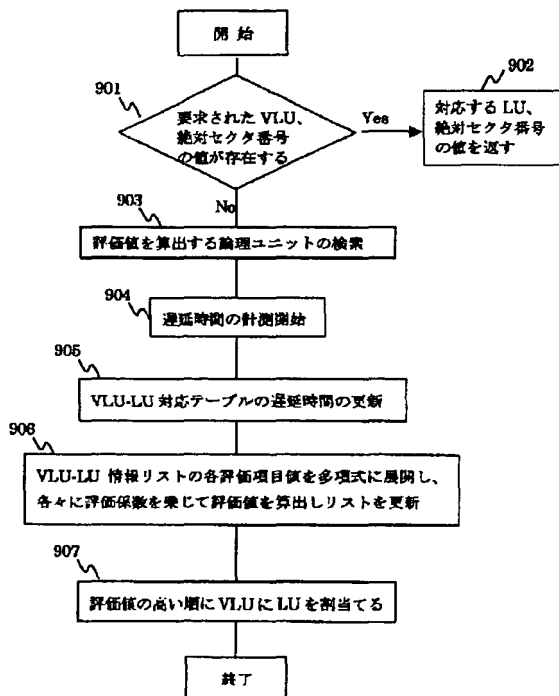
802

801

| VLU/LU 絶対セクタ番号(MB) | LUN 絶対セクタ番号(MB) | 遅延時間 (ms) | アクセス頻度(1/日) | バッチアクセス頻度(1/年) | 書き有無 | 評価値 | | |
|--------------------|-----------------|-----------|-------------|----------------|------|-----|---|------|
| 1 | 0- | 1 | 0- | 67 | 324 | 365 | 0 | 923 |
| 1 | 100- | 2 | 100- | 42 | 163 | 365 | 0 | 1106 |
| 1 | 200- | 3 | 0- | 147 | 0 | 52 | 0 | 1036 |
| 2 | 0- | 11 | 200- | 174 | 122 | 52 | 1 | 1216 |
| 2 | 100- | 12 | 0- | 33 | 273 | 52 | 1 | 1337 |
| 0 | 0- | 6 | 100- | 67 | 142 | 52 | 0 | 1079 |

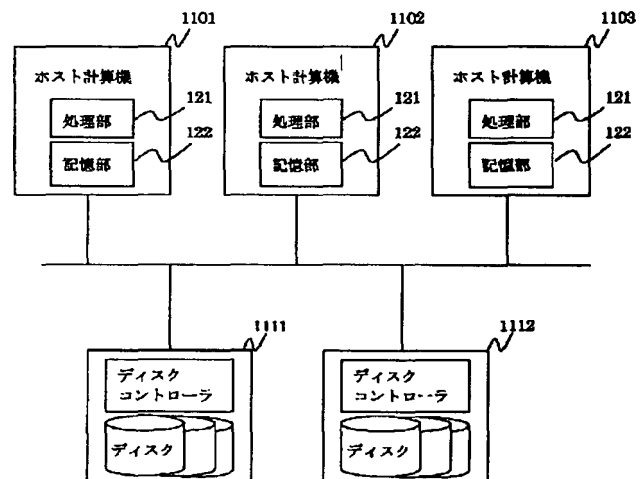
【図9】

図9



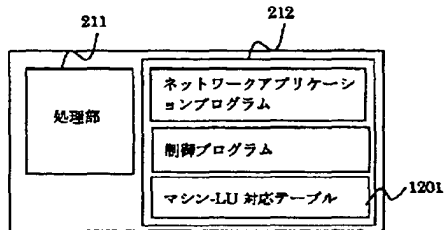
【図11】

図11



【図12】

図12



【図13】

図13

| 1301 | 1302 | 1303 | 1304 | 1305 | 1306 | 1307 | 1308 | 1309 | 1310 |
|--------|-------------|------|-------------|----------|---------------|---------------|-------------|------|------|
| マシン識別子 | 絶対セクタ番号(MB) | LUN | 絶対セクタ番号(MB) | 遅延時間(ms) | アクセス頻度(I/O/日) | バックアップ頻度(1/年) | ボリューム二重書き有無 | | 評価値 |
| 1 | 0- | 1 | 300- | 57 | 324 | 52 | 0 | | 923 |
| 1 | 100- | 2 | 100- | 57 | 163 | 365 | 0 | | 1216 |
| 1 | 200- | 3 | 200- | 57 | 163 | 365 | 0 | | 1216 |
| 2 | 0- | 4 | 300- | 57 | 163 | 365 | 0 | | 1216 |
| 2 | 100- | 5 | 400- | 57 | 163 | 365 | 0 | | 1216 |
| 0 | 0- | 6 | 200- | 57 | 0 | 52 | 0 | | 1379 |
| 0 | 0- | 7 | 300- | 57 | 0 | 52 | 0 | | 1379 |

【図14】

図14

